Author

Ali Razmara delights in the intellectual acrobatics research requires. As he says, "While searching for one piece of the puzzle, a far more important piece is revealed, leading to unexpected twists and turns and…a more critical examination of the task at hand." The project presented here required Razmara to use his scientific, technical and theoretical knowledge, something he did not expect when he first started doing research, as his original project was an investigation of a technical aspect of protein folding. Razmara's ultimate goal is to become "not only a resilient medical researcher, but also a benevolent physician who truly cares for the welfare of his patients."

Key Terms

- Magnetic Relaxation Dispersion (MRD)
- Molecular Dynamics (MD) Computer Simulation
- Molten Globule (MG)
- Protein Hydration

Hydration Dynamics of a Partially Unfolded Protein

Ali Razmara
Biological Sciences and Chemisty

Abstract

There has been substantial interest from both experimentalists and theoreticians in characterizing molten globule (MG) states of proteins. The almost-folded MG state is an ensemble of compact structures with persistent secondary structure, but disrupted tertiary interactions. Recently, ^{17}O magnetic relaxation dispersion (MRD) experiments by Denisov *et al.* probing protein hydration dynamics have suggested that the MG state is more structured and less hydrated than previously believed. This contradicts the traditional view that the MG is an expanded, loosely-packed protein with significant solvent exposure. In our project, molecular dynamics (MD) computer simulations have sought to reconcile the traditional picture of the MG state with the MRD data. Theoretical simulations were used to generate a plausible model of a partially-unfolded MG state of the barnase ribonuclease protein. Computed MRD observables from the MD trajectories strongly agreed with the experimental data. Room temperature analysis of hydrating water dynamics on the surface of the MG and native (N) states indicates faster water rotational dynamics for the MG. Our hypothesis is that water molecules reorient faster on the MG surface due to the loss of specific hydration sites that hinder their rotation in the N state. The experimental MRD data have been reinterpreted to support the traditional view that a large influx of water molecules actually penetrates the MG state upon volume expansion.

Faculty Mentor

Water, life's solvent, plays two vital roles in the function of biological macromolecules such as the enzymes that carry out biochemical reactions. One is to stabilize the specific structures that these molecules maintain in their functioning, native states. Another is that fast movement of water molecules promotes the flexibility of biological molecules required for their function. The paper by Ali Razmara reports a computer simulation study of water dynamics on the surface of a protein in its native state, and in a partially unraveled, denatured state. The results of Ali's research have led to the reconciliation of apparently contradictory experimental data on the structure and hydration of denatured proteins. The value of undergraduate research in a student's scholarly development cannot be understated. Creative thinking is not always necessary for success in the classroom, but it is essential for realizing our full intellectual potential. Research is an ideal experience for putting our textbook training to practical use, and for expanding our capacity for creative thinking.

Douglas Tobias
School of Physical Sciences

Introduction

Proteins are instrumental in many of the important biological functions of an organism, including enzymatic catalysis, mechanical support, molecular signaling, storage, transport, hormonal regulation, and defense against foreign substances. A functional protein is not merely a polypeptide, or a polymer of amino acids linked by peptide bonds, but also requires a precisely twisted, folded and globular shape. The perplexing phenomenon of protein folding is a spontaneous self-assembly process because the prerequisite information is encoded in the protein's amino acid sequence. During the protein folding process, the information contained within the linear amino acid sequence is translated into a three-dimensional, folded protein by an unresolved mechanism. The particular sequence of amino acids dictated by the gene uniquely determines the three-dimensional protein conformation, or shape, which governs how the protein will ultimately function.

As a cell synthesizes a polypeptide, the chain of amino acids folds spontaneously to assume the functional conformation of the protein. Hence, protein function is an emergent property resulting from the delicate molecular ordering of the 20 different amino acid building blocks. Three superimposed hierarchical levels of structure (primary, secondary and tertiary) comprise the complex architecture of a protein. The primary structure describes the protein's unique linear amino acid sequence encoded by the gene. Resulting from regular hydrogen bonding, the secondary structure refers to localized folded configurations of the primary amino acid sequence into delicate α–helix coils and β–pleated sheets. Superimposed on the patterns of secondary structure is the tertiary structure of a protein, which encompasses the global three-dimensional shape into which the entire polypeptide chain is eventually folded. This unique conformation gives a protein its specific function.

It seems possible to ascertain the mechanism of protein folding by correlating the protein's linear primary structure and tertiary conformation, but this biological phenomenon is not as simple as one may believe. Along the folding pathway, many proteins progress through several intermediate structures before reaching their respective native functioning states. Understanding this process is as important as it has proven challenging. Most notably, knowledge of the protein-folding mechanism on a molecular level will provide deep insight into various human diseases caused by proteins that have malfunctioned. Likewise, construction and design of artificial proteins capable of performing specific biological tasks will aid in the development of drug design technology in the chemical and pharmaceutical industries. In addition, scientists and physicians will be able to make sense of enormous amounts of data from various genome projects, which will impact our understanding of the rules governing the important biological functions of proteins.

A full understanding of the mechanism of protein folding requires knowledge of the structures, relative energetics, and dynamics of ensembles of intermediate species populating the folding pathway. Thus, considerable interest has been expressed by both experimentalists and theoreticians in characterizing compact denatured or molten globule (MG) states of proteins (Brooks III, 1998; Dobson, 1992; Onuchic, 1995; Ptitsyn, 1995; Shortle, 1993). Because these MG states are disordered, a detailed picture is lacking. Several experimental techniques have indirectly investigated these partially-unfolded intermediate states. In particular, nuclear magnetic resonance (NMR) and optical spectroscopy have provided glimpses of residual structure and interactions in denatured states (Eliezer, 1998; Hughson et al., 1990; Kuwajima, 1976; Schulman et al., 1997), while global views of the overall sizes and shapes of denatured proteins have been obtained by small angle neutron and X-ray scattering (Calmetter et al., 1993; Kataoka, 1997; Lattman, 1994). Structurally, the MG state is usually described as an ensemble of compact structures with persistent secondary structure but disrupted tertiary interactions and extensive solvent penetration of the internal hydrophobic cores. This has been suggested by calorimetry measurements (Freire, 1995).

Traditionally, the MG state is viewed as an expanded, solvent-penetrated, and late-protein-folding intermediate distinguished from the native (N) state. This expansion of the protein volume has been perceived to imply that hundreds of water molecules actually penetrate the MG state (Fink, 1995; Kharakoz and Bychkova, 1997; Uchiyama et al., 1995). Consequently, MG models evoke a sizable degree of internal hydration of the protein macromolecule (Finkelstein and Shakhnovich, 1989; Williams et al., 1997). Recently, the ^{17}O magnetic relaxation dispersion (MRD) (Halle et al., 1999) experimental technique has been utilized to provide detailed dynamic information regarding the external hydration of proteins in water solutions. The highlight of the MRD technique stems from the fact that it provides more direct information regarding the order, extent and dynamics of the hydrating water molecules than previous experiments (Denisov et al., 1999).

These recent NMR relaxation experiments have sought to

characterize the structures of MG states by probing the dynamics of internal and surface bound water molecules. The MRD data provides a global measure of external protein hydration in the form of the experimentally observable quantity $N_S\rho_S$. The extent of hydration is defined by N_S, which is the average number of surface bound water molecules. The dynamics of hydration are defined by the relative dynamic retardation, ρ_S, which is the slow-down factor associated with water molecules next to the protein surface. MRD measurements (Denisov et al. 1999) on the hydration of the native and MG states of several proteins have suggested a controversial picture of the MG state that is quite different from the prevailing view of an expanded, loosely-packed MG protein with significantly greater solvent exposure than the N state (Shortle, 1999). Among other things, Denisov *et al.* found that the quantity $N_S\rho_S$ was essentially the same in the N and MG states of a given protein (1999). Thus, they concluded that the external hydration of the investigated MG protein differed very little from that of the corresponding N protein, both in its extent (N_S) and in its dynamics (ρ_S). These new experiments challenge the traditional picture of the MG state as expanded and solvent-penetrated compared to the N state (Fink, 1995).

Computer simulations of classic dynamics of atoms and molecules have led to valuable insight and improved understanding in diverse areas of science, including materials design, drug discovery, and protein structure. Computer simulation of the dynamics of bio-molecules by the molecular dynamics (MD) technique (Brooks et al., 1988) yields the possibility of describing and understanding the structure-dynamics-function relationships of molecular processes such as protein folding changes in terms of interactions at the atomic level. MD computer simulation is the science of simulating the motions of a system of particles by solving equations of motion given a force field, and thereby generating detailed atomic trajectories. The time evolution of the positions of the atoms of the system can provide the ultimate microscopic picture, which may serve to explain observed macroscopic behavior of a molecular system. Once the reliability of the molecular models, force fields, and computational procedures has been established by comparison of simulated properties with experimental data, computer simulation can be a very powerful tool to predict molecular properties that are inaccessible to experimental probes. Furthermore, the advantage of using MD simulations to examine protein denaturation lies in knowing the protein structure at every step. Using these simulations also presents an opportunity to gain detailed insight and knowledge into different factors that influence the stability of proteins in their native states. Thus, the microscopic details derived from these theoretical studies based on atomic scale MD simulations provide models for elucidation of experimental data and novel insight into the protein folding mechanism. In this research, MD computer simulation studies were performed using the 110 amino acid barnase ribonuclease protein, a convenient system for studying protein folding. These theoretical studies were aimed at reconciling the apparently contradictory view of the MG state derived from MRD data through detailed analysis of the hydration of the N state and a reasonable model for the MG state. By understanding the hydration dynamics of these nearly-folded MG proteins, it is conceivable to gain further insight into the rules governing protein folding and why native proteins are functional and stable. A fundamental understanding of the way proteins fold and function may pave the way for future scientists to alleviate human suffering in many diseases by bioengineering newly designed and functional proteins that will correct proteins that have malfunctioned.

Materials and Methods

Magnetic Relaxation Dispersion Experimental Background

Before we present the results of our calculations, it is first instructive to consider the NMR relaxation measurements and the interpretation of them in order to demonstrate that the experiments provide dynamics data on a time scale that is sampled well by atomistic MD computer simulations. In particular, ^{17}O relaxation was considered, which reports exclusively on water dynamics. In the case of fast exchange between different microenvironments, there are three possible microenvironments (bulk (B), surface (S), and internal (I), shown in Figure 1), each having different average relaxation, or rotational correlation times, τ. While the most internal water molecules are trapped and rotate the slowest with $\tau_I > 1$ nsec, the bulk water molecules rotate the fastest with $\tau_B \sim 3$ psec. The rotational dynamics of the surface bound water molecules ($\tau_S \sim 10$ psec) provide useful information about protein structure.

In an MRD experiment, the quantity $N_S\rho_S$ is determined from the frequency dependence of the longitudinal relaxation rate. Under mild assumptions, spin-lattice relaxation theory may be used to relate the dynamical retardation factor, ρ_S, to a ratio of rotational correlation times, *i.e.* $\rho_S = (\tau_S / \tau_B) - 1$. For proteins of known structure, N_S may be estimated from the solvent-accessible surface area, while τ_B can be measured independently. Thus, a measurement $N_S\rho_S$ can provide an estimate of τ_S for native proteins of known structure. Denisov *et al.* have measured $\rho_S \approx 4$, which gives $\tau_S \approx$

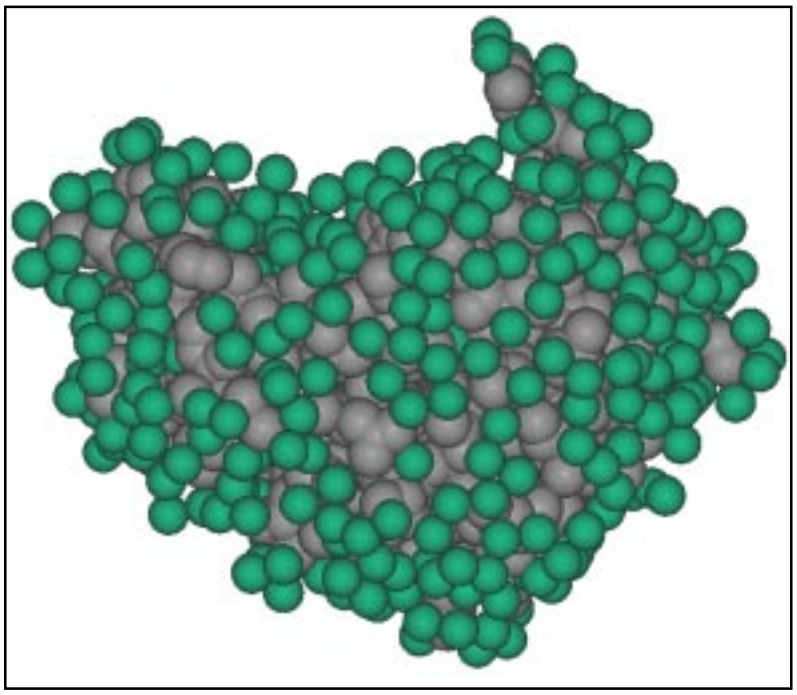

Figure 1

Space-filling model of the native barnase protein (gray spheres) in solution showing surface bound water molecules (green spheres). There are also internal trapped water molecules (not visible) buried in cavities as well as bulk water molecules (not shown) that are not bound to the protein surface. The rotational dynamics of the surface bound water molecules provide the most useful information about the structure of the protein. These surface bound water molecules also rotate on a time scale that is well sampled by MD computer simulations.

15 psec for several native globular proteins (1999). This is a time scale that is well-sampled by MD simulations. A practical consequence is that we can use atomistic MD computer simulations to make contact with the MRD measurements.

Several quantities were calculated in order to compare with the MRD experimental results. For the calculation of ρ_S, we needed rotational correlation times for the surface and bulk water molecules. We obtained these from the second-rank rotational correlation functions:

(1)

where $\hat{\mathbf{u}}$ is a molecule-fixed unit vector (the results presented here are for the H–H vector, but other choices, like the O–H dipole moment, give essentially identical results for ρ_S), and $P_2(x)$ is the second Legendre polynomial of argument x. These correlation functions do not exhibit an exponential decay, as expected for rotational Brownian motion, but rather are well-described by a stretched exponential, or Kohlrausch-Williams-Watts (KWW) form:

$$C(t) = \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \qquad (2)$$

The KWW form arises from spatial and/or temporal inhomogeniety in the motion (Rocchi et al., 1998). It is a signature of a distribution of relaxation times. For example, individual water molecules interacting with different parts of a rough, chemically heterogeneous protein surface are expected to have different reorientation rates. Although the rotational correlation function for each water molecule may exhibit an exponential decay with a characteristic correlation time, when all of the correlation functions with different correlation times are averaged to give the overall $C(t)$, the result is a stretched exponential (Abseher et al., 1996). For comparison with experiment, the average correlation time may be computed from (Settles and Doster, 1996):

$$\langle\tau\rangle = \int_0^\infty C(t)dt = \frac{\tau}{\beta}\Gamma\left(\frac{1}{\beta}\right) \qquad (3)$$

where τ and β are parameters from fits to the KWW form above. Subsequently, the dynamic retardation factor was calculated from:

$$\rho_S = \left(\langle\tau\rangle_S / \langle\tau\rangle_B\right) - 1 \qquad (4)$$

Molecular Dynamics Computer Simulation Background

The essential elements for MD simulations are the potential energy for the particles, from which forces can be calculated, and Newton's equations of motion governing the dynamics of the particles. The equations of motion are numerically integrated on a computer, producing a "trajectory" of positions and velocities at a series of discrete time steps, Δt. A typical molecular force field or effective potential for a system of N atoms with masses m_i (i = 1, 2, … , N) and Cartesian position vectors $\boldsymbol{r}_i$ has the form:

$$V(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \Sigma_{bonds}\, \tfrac{1}{2}K_b[b - b_0]^2 + \Sigma_{angles}\, \tfrac{1}{2}K_\theta[\theta - \theta_0]^2 + \Sigma_{improper\ dihedrals}\, \tfrac{1}{2}K_\xi[\xi - \xi_0]^2 + \Sigma_{dihedrals}\, K_\varphi[1 + \cos(n\varphi - \delta)] + \Sigma_{pairs\ (i,j)}[C_{12}(i,j)/r_{ij}^{12} - C_6(i,j)/r_{ij}^{6} + q_iq_j/(4\pi\varepsilon_0\varepsilon_r r_{ij})] \qquad (5)$$

where the first term is a harmonic potential in which the minimum energy bond length b_0 and the force constant K_b vary with the particular bond type and describes the covalent bond stretching interaction along bond length b. The second term represents the bond angle bending for a three-body

interaction in a similar harmonic form. As for the four-body dihedral angle interactions, two forms are used. The third term represents a harmonic potential for the improper dihedral angles ξ that are not allowed to make transitions. The fourth sinusoidal term describes the dihedral angles φ that are allowed to make 360° turns. As a sum over all pairs of atoms, the last sum term describes the effective non-bonded interaction, which is comprised of van der Waals and Coulombic interactions between atoms i and j with charges q_i and q_j at a distance r_{ij} (van Gunsteren and Berendsen, 1990).

A convenient system for studying protein folding has been the barnase ribonuclease protein (Fersht, 1993). Barnase is a 110 amino acid RNA endonuclease from the bacteria *Bacillus amyloliquefaciens*. It is a single-domain protein consisting of three α-helices followed by a five-stranded anti-parallel β-sheet and three hydrophobic internal cores (Figure 2) (Bycroft et al., 1991). Because of its small size and previous extensive experimental studies (Matouschek et al., 1992; Serrano et al., 1992), as well as available configurations from an investigation of some technical aspects of unfolding simulations, the barnase protein was used in our theoretical investigation.

Results and Discussion

Molecular Dynamics Computer Simulations of the Native and Molten Globule States

The initial coordinates of the native barnase protein X-ray crystal structure were obtained from the Protein Data Bank (PDB) at 1.5 Å resolution (PDB entry 1A2P). MD simulations were performed with the PINY_MD program (Tuckerman et al., 2000) and the CHARMM 22 force field (MacKerell et al., 1998) using the SPC/E deuterated water model representative of the solvent. Each run consisted of 10,000 steps (40 psec) with a time step of 4.0 fsec. The system simulated consisted of 12,385 atoms (108 amino acid residues, 2 chloride counter ions, and 3,561 water molecules).

Employing periodic boundary conditions and the Ewald sum, the native state was first simulated at room temperature (300 K) and compared to the crystal structure to check the force field. At the end of the (640 psec, constant pressure, and 300 K) simulation of the native protein state, the root-mean-square deviation (RMSD) from the starting crystal structure for the native (N) structure was 0.92 Å for the protein backbone α-carbons, whereas the RMSD for all non-hydrogen atoms was 1.29 Å. The radius of gyration (R_g), measuring the hydrodynamic radial size of the protein, of the crystal structure was 13.5 Å, while that of the N structure was a comparable 13.7 Å. These results demonstrate that the force field applied was indeed acceptable. Furthermore, other global parameters utilized include monitoring of the solvent-accessible surface area (SASA) using a 1.6 Å probe radius and calculation of the number of bound waters using a 4.0 Å cut-off radius.

In the subsequent MD simulation, the equilibrated native barnase protein was unfolded in a 1.2 nsec, constant tem-

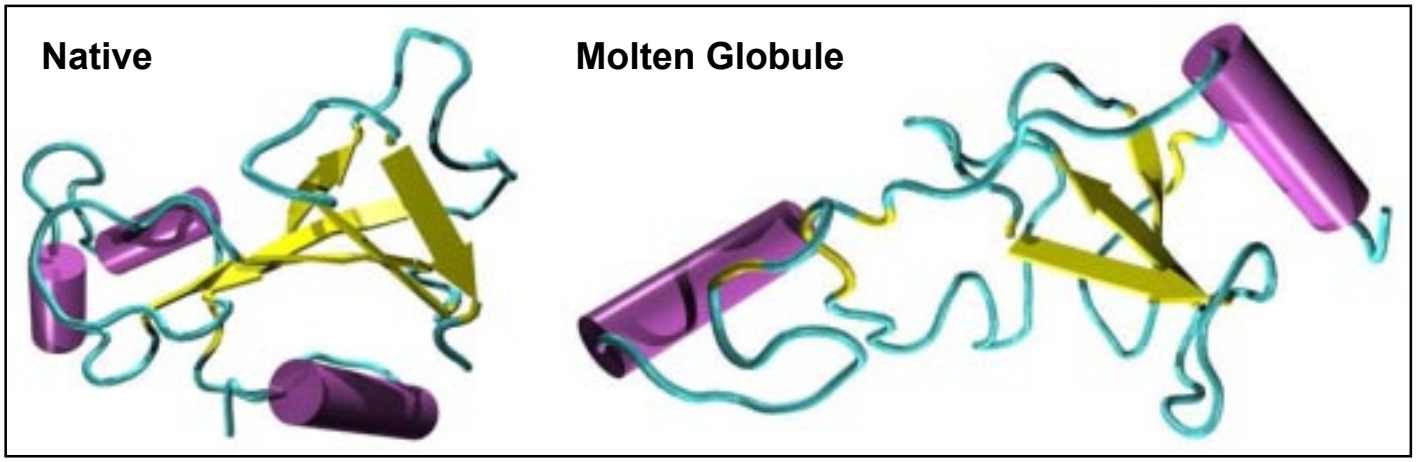

Figure 2

Cartoon diagram representations of the native (N) (left) and molten globule (MG) (right) states of barnase. The purple cylinders represent α-helices, the yellow arrows depict the β-sheet strands, and the cyan tubes are the loops and turns. The N state (left) is a globular protein containing three α-helices followed by a five-stranded anti-parallel β-sheet and three hydrophobic internal cores. After MD computer simulation at 500K and constant pressure, a partially unfolded MG (right) structure was generated. The MG protein lost its three-dimensional globular shape to become an elongated stable structure. This indicates significant internal hydrophobic core disruption, while largely retaining (~75%) the native α-helix and β-sheet secondary structural elements.

perature (500 K) and atmospheric pressure run. A partially unfolded structure was used to initiate a simulation at 300 K. Various structural properties of the N and MG states computed from our simulations are compared in Table 1. The partially unfolded structure did not significantly deviate over several hundred picoseconds, maintaining a roughly 8 Å protein backbone α-carbon RMSD from the crystal structure, 76% of the native secondary structural hydrogen bonds, and a 14% increase in the R_g compared to the native structure. Since the extent of expansion relative to the native state was in the range observed for real MG states (Kataoka, 1997; Smith et al., 1999; Sosnick and Trewhella, 1992), our artificially generated, partially unfolded barnase structure was regarded as a putative member of an MG ensemble for the purpose of comparison to the NMR relaxation experiment.

Table 1

Parameters comparing crystal, native (N), and molten globule (MG) structures of barnase.

Structure	*α–Carbon RMSD (Å)*	*All non-H atom RMSD (Å)*	*R_g (Å)*	*% Native 2° Structure*	*#Bound Waters*
Crystal	-	-	13.5	-	-
N	0.92	1.29	13.7	-	430
MG	8.32	8.84	15.6	76	548

Figure 2 shows cartoon diagrams of the simulated native and molten globule states of barnase. Dramatic unfolding changes were immediately evident from the distinguishing structural characteristics of the two structures. Although the α-helices remained mostly stable, the anti-parallel β-sheet elements were disrupted. More importantly, the main internal hydrophobic core was completely disrupted, indicating the loss of globular three-dimensional shape in the MG protein. Thus, the computer generated MG structure of barnase was considerably expanded having disrupted global tertiary structure but largely retained its native local secondary structure. Having established a reasonable model for the MG state, we now proceed to make contact with the MRD experimental results.

Another indicator that clearly distinguished the MG and N structures was the directly calculated number of bound waters (Figure 1). Using our simulation analysis, we were able to accurately count every bound water molecule on the surface of the different protein states. Defining the number of surface water molecules, N_S, as the number of water molecules within a certain distance (in this case, 4 Å) from any protein atom (Rocchi et al., 1997), averaged over the trajectory, directly gave 430 surface water molecules for the N state, and 548 for the MG state. Thus, we found that transition from the N to the MG state was accompanied by a substantial increase (≈ 100 waters, or ≈ 25% for this 12 kDa barnase protein) in the number of water molecules in contact with the protein surface, consistent with the traditional picture of the molten globule.

Hydration Dynamics of the Native and Molten Globule States

Next we considered estimates of the quantity $N_S \rho_S$ from our simulations to check for consistency with the MRD data. The correlation functions, $C(t)$, computed according to equation (1) from our simulations of bulk D_2O and D_2O in contact with the surface of the N and MG states of barnase are shown in Figure 3. In the protein systems, the contribution of a given bound water to the $C(t)$ was included only while that water was near the protein surface (in our case, within 4 Å). Consistent with previous work (Abseher et al., 1996; Rocchi et al., 1998), the data in Figure 3 showed that the rotational motion of water molecules near the protein was significantly slowed relative to the bulk. Moreover, the data demonstrated that, on average, the water molecules near the N state reoriented slightly slower than those near the MG state. The $C(t)$ shown were averages over many water molecules and time origins. Therefore, we believe the slight difference in $C(t)$ between the native and the single initial configuration of the MG state considered is significant.

From fits of equation (2) to the data in Figure 3, using equation (3) we obtained $\langle \tau \rangle_B$ = 2.4 psec for bulk D_2O, and $\langle \tau \rangle_S$ = 14.4 psec and 11.1 psec for surface D_2O in the N and MG solutions, respectively. The bulk result was in reasonable agreement with the experimental value of 3.1 psec (Halle and Wennerstrom, 1981), and the relative dynamic retardation factor, ρ_S = 4.9 for the N state, was in the range of values extracted from MRD data on native proteins (Denisov and Halle, 1995; Halle, 1999). For the MG state, we obtained ρ_S = 3.6, and hence we found (using the average numbers of surface water molecules reported above for N_S), in agreement with the MRD results, that the quantity $N_S \rho_S$ was essentially the same in the N and MG states, *i.e.* $N_S \rho_S$ (MG) / $N_S \rho_S$ (N) = 1.07. Table 2 summarizes the surface water rotational dynamics quantities for the N and MG states of barnase.

Our results suggest that the quantity $N_S \rho_S$ remains essentially unchanged on passing from the N to the MG state because a

Table 2

Surface bound and bulk water rotational dynamics quantities

Water Type	$<\tau>$	ρ_S	N_S	$N_S \rho_S$
Bulk	2.4	-	-	-
N bound	14.4	4.9	430	2107
MG bound	11.1	3.6	548	1973

Table 3

Protein surface water hydrogen bond calculation data

Structure	*Bound Water HB*		*Avg. HB ****	τ_{HB}
	*P-W HB **	*W-W HB ***		
N	0.48	2.35	205	0.52
MG	0.43	2.39	235	0.49

**Average number of Protein-Water Hydrogen Bonds per water molecule*

***Average number Water-Water Hydrogen Bonds per water molecule*

****Average total number of Protein-Water Hydrogen Bonds*

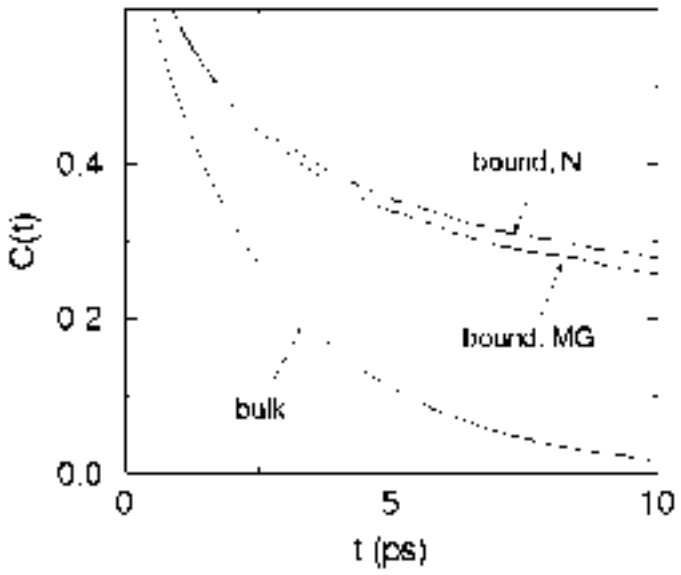

Figure 3

Second-rank rotational correlation functions, *C(t)*, computed from MD simulations of bulk D_2O and D_2O close ("bound") to the surface of barnase in the N state and a model of the MG state at 27°C. The *C(t)* for bulk waters decayed (dropped off) faster than the *C(t)* for bound waters, meaning bulk waters reoriented faster than bound waters. *C(t)* for bound MG waters decayed faster than that of bound N waters, indicating bound MG waters reoriented faster than bound N waters. Although this difference in correlation function decay for bound MG and bound N waters seemed to be small, it was a significant difference when mean correlation times were calculated.

substantial increase (27%) in N_S is accompanied by a substantial decrease (36%) in ρ_S; in other words, the water molecules are reorienting significantly faster, on average, in the vicinity of the MG state. Utilizing MD computer simulation studies, we have reconciled the traditional picture of the MG state as an expanded and solvent penetrated partially unfolded intermediate compared to the native state with the MRD experimental data, which has been taken to support no difference in the hydration of the native and MG states. From our simulations we found that 60% of the additional solvent-accessible surface area exposed in the MG state was occupied by hydrophobic side chains. A simulation of BPTI suggested that water residence times at hydrophobic residues were shorter than at hydrophilic residues (Muegge and Knapp, 1995). However, recently a simulation of Azurin asserts that comparison of water residence times between different proteins is complicated by different factors such as the chemical environment, surface shape, and solvent-accessibility (Luise et al., 2000). Indeed, our analysis of residence time distributions for both the N and MG states of barnase revealed a shift towards shorter residence times for the MG state.

We originally hypothesized that the faster rotation of the surface water molecules might have to do with the additional exposure of hydrophobic, non-polar groups in the MG state. Several analytical methods were employed to test and revise our original hypothesis. First, analysis of the correlation functions for specific hydrophobic, polar and charged amino acid residue types that were solvent exposed either commonly or uniquely in the N and MG states of barnase unexpectedly provided little differences. Second, calculation of the N and MG surface physical curvature elements did not differentiate the different water dynamics on the surface of these protein states. Lastly, surface water hydrogen bond (HB) calculations (Table 3) offered an explanation of the differences in the surface water dynamics. On average, the MG had more surface hydrogen bonds because of a greater number of bound water molecules, but on average each of these water molecules had shorter lifetimes (τ_{HB}), referring to the average time it takes for a protein-water hydrogen bond to break. Next, examining the bound water molecules in both structures, the transition from the N to MG was accompanied by a shift from longer-lived protein-water hydrogen bonds (P-W HB) towards shorter-lived water-water hydrogen bonds (W-W HB). Our revised hypothesis is that water molecules reorient faster next to the surface of the MG not due to proximity to the "slippery" hydrophobic groups, but because of the loss of particular hydration sites that hinder free rotation of the surface waters in the N state. This causes the rotational cor-

relation times to shift toward smaller values in the MG state, resulting in smaller relative dynamic retardation, *i.e.* ρ_S *(MG)* < ρ_S *(N)*.

Conclusion

By employing MD computer simulations, a reasonable model for the MG was generated and subsequently compared to a controversial NMR experiment. The experimenters (Denisov et al., 1999) interpreted the constancy of the experimental quantity $N_S\rho_S$ from the MRD data to imply that nonnative proteins, especially MG intermediates, were more structured and less hydrated than previously believed. They assumed that the water molecules "see" the same environment in the N and MG states, and therefore that there was no change in their inherent rotation rate, *i.e.* ρ_S *(N)* ~ ρ_S *(MG)*. Thus, for the quantity $N_S\rho_S$ to be constant, they concluded that the number of surface bound water molecules did not change, *i.e.* N_S *(N)* ~ N_S *(MG)*, presenting an apparent paradox that is at odds with previous X-ray scattering and calorimetry studies, which advance the traditional view of the MG state as more hydrated and expanded than the N state. In other words, since the quantity $N_S\rho_S$ was observed to be constant, these experimenters assumed the slow-down factor to be similar, *i.e.* ρ_S *(N)* ~ ρ_S *(MG)*, in the transition from the N to MG state. This led to their incorrect conclusion that the number of MG surface bound waters did not change relative the N state, *i.e.* N_S *(N)* ~ N_S *(MG)*, thereby directly challenging the usual view of an expanded solvent penetrated MG state.

In contrast to this recent MRD experiment (Denisov et al., 1999), without any assumptions, we were able to separately calculate both the components of the experimentally derived quantity $N_S\rho_S$. From the atomic simulated trajectories, the surface number of bound waters, N_S, was directly calculated with no assumptions. Also, the correlation functions were utilized to find the rotational times, and therefore the dynamic retardation factor, ρ_S, was calculated. Moreover, our MD theoretical investigations demonstrate that the surface waters actually experience a different rotational environment, *i.e.* ρ_S *(MG)* < ρ_S *(N)*, along with an increase in the number of surface bound waters for the MG, *i.e.* N_S *(MG)* > N_S *(N)*, thus the quantity $N_S\rho_S$ is constant. Our computer simulation results agree with the NMR relaxation experimental results, but our interpretation of the data distinguishes our analysis from that of these experimenters (Denisov et al., 1999). Thus, by analyzing the hydration dynamics of this MG, we are able to reconcile the traditional view of the MG state with the MRD experiment. Our results indicate that the increase in the number of surface bound waters is balanced by a decrease in the relative dynamic retardation factor in the transition of the N to MG state. Our conclusions support the traditional notion that a large influx of water molecules actually penetrates the MG state upon volume expansion, therefore distinguishing the partially unfolded MG intermediate from the native state.

Our theoretical investigations have simulated a water dynamics experiment on the surfaces of both native, functioning and almost folded, nonfunctioning protein states. By understanding the stability of these partially unfolded molten globule intermediates of proteins, we can gain insight into the rules that govern the important mechanism of protein folding. The implications of such fundamental knowledge on the molecular level will lead to the future development of artificially designed but biochemically active proteins, replacing the important task of many malfunctioned proteins underlying various human diseases.

Acknowledgements

This project was a collaborative team effort for which I am grateful. I would like to thank my mentor, Dr. Douglas Tobias, for his guidance, inspiration, and motivation. I would also like to thank William Kuo for his valuable contributions. Furthermore, I am appreciative of monetary support by the UROP/PUF grant and the Arnold and Mabel Beckman Scholars Program.

Works Cited

Abseher, R., H. Schreiber, and O. Steinhauser. "The Influence of a Protein on Water Dynamics in its Vicinity Investigated by Molecular Dynamics Simulation." Proteins: Structure, Function, and Genetics 25 (1996): 366-378.

Brooks, C.L., M. Karplus, and B.M. Pettitt. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. New York: Wiley, 1988.

Brooks III, C.L. "Simulations of Protein Folding and Unfolding." Current Opinion in Structural Biology 8 (1998): 222-226.

Bycroft, M., et al. "Determination of the Three-Dimensional Structure of Barnase Using Nuclear Magnetic Resonance Spectroscopy." Biochemistry 30 (1991): 8697-8701.

Calmetter, P., et al. "Configurational Distribution of Denatured PhosPhoglycerate Kinase." Journal of Molecular Biology 231 (1993): 840-848.

Denisov, D.P., B.H. Jonsson, and B. Halle. "Hydration of Denatured and Molten Globule Proteins." Nature Structural Biology 6 (1999): 253-260.

Denisov, V.P., and B. Halle. "Protein Hydration Dynamics in Aqueous Solution. A Comparison of Bovine Pancreatic Trypsin Inhibitor and Ubiquitin by Oxygen-17 Spin Relaxation Dispersion." Journal of Molecular Biology 245 (1995): 682-697.

Dobson, C.M. "Unfolded Proteins, Compact States, and Molten Globules." Current Opinion in Structural Biology 2 (1992): 6-12.

Eliezer, D., et al. "Structural and Dynamic Characterization of Partially Folded States of Apomyoglobin and Implications for Protein Folding." Nature Structural Biology 5 (1998): 148-155.

Fersht, A.R. "Protein Folding and Stability: The Pathway of Folding of Barnase." FEBS Letters 325 (1993): 5-16.

Fink, A.L. "Compact Intermediates States in Protein Folding." Annual Review of Biophysics and Biomolecular Structure 24 (1995): 495-522.

Finkelstein, A.V., and E.I. Shakhnovich. "Theory of Cooperative Transitions in Protein Molecules. II. Phase Diagram for a Protein Molecule in Solution." Biopolymers 28 (1989): 1681-1694.

Freire, E. "Thermodynamics of Partially Folded Intermediates in Proteins." Annual Review of Biophysics and Biomolecular Structure 24 (1995): 141-165.

Halle, B. "Water in Biological Systems: The NMR Picture." Hydration Processes in Biology: Theoretical and Experimental Approaches. Ed. M.C. Bellissent-Funel. Amsterdam: IOS Press (1999): 233-249.

Halle, B., V.P. Denisov, and K. Venu. Modern Techniques in Protein NMR. Ed. L.J. and Krishna Berliner, N.R. Vol. 17. New York: Plenum Press, 1999.

Halle, B., and H. Wennerstrom. "Interpretation of Magnetic Resonance Data from Water Nuclei in Heterogeneous Systems." Journal of Chemical Physics 75 (1981): 1928-1943.

Hughson, F.M., P.E. Wright, and R.L. Baldwin. "Structural Characterization of a Partly Folded Apomyoglobin Intermediate." Science 249 (1990): 1544-1548.

Kataoka, M. "Structural Characterization of the Molten Globule of α-lactalbumin by Solution X-ray Scattering." Protein Science 6 (1997): 422-430.

Kharakoz, D.P., and V.E. Bychkova. "Molten Globule of Human α-lactalbumin: Hydration, Density, and Compressibility of the Interior." Biochemistry 36 (1997): 1882-1890.

Kuwajima, K. "Three-State Denaturation of α-lactalbumin by Guanidium Hydrochloride." Journal of Molecular Biology 106 (1976): 266-280.

Lattman, E.E. "Small Angle Scattering Studies of Protein Folding." Current Opinion in Structural Biology 4 (1994): 87-92.

Luise, A., M. Falconi, and A. Desideri. "Molecular Dynamics Simulation of Solvated Azurin: Correlation Between Surface Solvent Accessibility and Water Residence Times." Proteins: Structure, Function, and Genetics 39 (2000): 56-67.

MacKerell, A.D., et al. "All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins." Journal of Physical Chemistry B 102 (1998): 3586-3616.

Matouschek, A., et al. "The Folding of an Enzyme V. H/2H Exchange-Nuclear Magnetic Resonance Studies on the Folding Pathway of Barnase: Complementarity to and Agreement with Protein Engineering Studies." Journal of Molecular Biology 224 (1992): 837-845.

Muegge, I., and E.W. Knapp. "Residence Times and Lateral Diffusion of Water at Protein Surfaces: Application to BPTI." Journal of Physical Chemistry 99 (1995): 1371-1374.

Onuchic, J.N. "Toward an Outline of the Topography of a Realistic Protein-Folding Funnel." Proceedings of the National Academy of Sciences, USA 92 (1995): 3626-3630.

Ptitsyn, O.B. "Molten Globule and Protein Folding." Advances in Protein Chemistry 47 (1995): 83-229.

Rocchi, C., A.R. Bizzarri, and S. Cannistraro. "Water Dynamical Anomalies Evidenced by Molecular-Dynamics Simulations at the Solvent-Protein Interface." Physical Review E 57 (1998): 3315-3325.

Rocchi, C., A.R. Bizzarri, and S. Cannistraro. "Water Residence Times Around Copper Plastocyanin: A Molecular Dynamics Simulation Approach." Chemical Physics 214 (1997): 261-276.

Schulman, B., et al. "A Residue Specific NMR View of the Non-Cooperative Unfolding of a Molten Globule." Nature Structural Biology 4 (1997): 630-634.

Serrano, L., A. Matouschek, and A.R. Fersht. "The Folding of an Enzyme III. Structure of the Transition State for Unfolding of Barnase Analyzed by a Protein Engineering Procedure." Journal of Molecular Biology 224 (1992): 805-818.

Settles, M., and W. Doster. "Anomalous Diffusion of Adsorbed Water: A Neutron Scattering Study of Hydrated Myoglobin." Faraday Discussions 103 (1996): 269-279.

Shortle, D. "Denatured States of Proteins and their Roles in Protein Folding and Stability." Current Opinion in Structural Biology 3 (1993): 66-74.

Shortle, D. "Protein Folding as Seen from Water's Perspective." Nature Structural Biology 6 (1999): 203-205.

Smith, L.J., C.M. Dobson, and W.F. van Gunstern. "Side-Chain Conformational Disorder in a Molten Globule: Molecular Dynamics Simulations of the α-state of Human α-lactalbumin." Journal of Molecular Biology 286 (1999): 1567-1580.

Sosnick, T.T., and J. Trewhella. "Denatured States of Ribonuclease A Have Compact Dimensions and Residual Secondary Structure." Biochemistry 31 (1992): 8329-8335.

Tuckerman, M.E., et al. "Exploiting Multiple Levels of Parallelism in Molecular Dynamics Based Calculations via Modern Techniques and Software Paradigms on Distributed Memory Computers." Computer Physics Communications 128 (2000): 333-376.

Uchiyama, H., et al. "Effects of Amino Acid Substitutions in the Hydrophobic Core of α-lactalbumin on the Stability of the Molten Globule State." Protein Engineering 8 (1995): 1153-1161.

van Gunsteren, W.F., and H.J.C. Berendsen. "Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry." Angewandte Chemie International Edition 29 (1990): 992-1023.

Williams, M.A., J.M. Thornton, and J.M. Goodfellow. "Modeling Protein Unfolding: Hen Egg-White Lysozyme." Protein Engineering 10 (1997): 895-903.